A Multidisciplinary Approach to Enhancing Infantry Training through Immersive Technologies

Fournier, H.; Lapointe, J.-F.; Munteanu, C.; Emond, B.; Kondratova, I.

A Multidisciplinary Approach to Enhancing Infantry Training through Immersive Technologies

Fournier H., Lapointe, J.-F., Munteanu C., Emond B., & Kondratova, I.
Institute for Information Technology, National Research Council of Canada
[Helene.Fournier; Jean-Francois.Lapointe; Cosmin.Munteanu; Bruno.Emond; Irina.Kondratova]@nrc-cnrc.gc.ca

ABSTRACT

There is a growing interest in the Canadian Army for using off-the-shelf computer games in training because of the interactivity and engagement they create for the trainees. Unfortunately, existing solutions do not always satisfy the aggregated training needs of infantry. In this paper, we briefly describe the challenges and shortcomings of both physical and virtual training systems, and the specific training needs identified through field studies in partnership with the Department of National Defence Canada. We then describe research solutions addressing these specific challenges, such as novel interaction methods in immersive environments, usable speech recognition, simulated weapons (laser rifles, electronic flashbang, etc.), flexible serious gaming platforms with intelligent agents and cognitive models, and mobile control interface for instructors. We also present the outcomes of field observations revealing what additions are required to improve learning scenarios for serious games.

Within our research project we created a platform for developing and validating novel interaction methods, technologies, and devices that create mixed-reality immersive training systems that are safer, more cost efficient, and more effective for learning than physical environments or purely virtual reality systems. The immediate use of our mixed-reality system is in practicing engagement skills and training personnel in the rapid application of judgment and of rules of engagement and in the use of force.

We conclude by describing the findings of human-subject evaluations conducted with an implementation of our research platform, presenting lessons from technology-specific feasibility, to educational/learning impact, and to human factors of interacting with an intelligent, immersive, training system. These findings provide encouraging evidence that such solutions will allow infantry, law enforcement, or public safety personnel to train using a virtual environment in a manner similar to training in a physical environment.

ABOUT THE AUTHORS

Helene Fournier: Dr. Helene Fournier is a Research Officer with the Learning and Collaborative Technologies research group at the National Research Council Institute for Information Technology (NRC-IIT). Dr. Fournier joined the NRC in 2002 and holds a Ph.D. in Educational Psychology from McGill University. Her research area is education and technology and she has participated in several research projects focused on the application and evaluation of advanced technologies in the training sector, in distance education, and more recently in learner-centered research and development of a Personal Learning Environment.

Jean-Francois Lapointe: Dr. Lapointe joined NRC-IIT as a Research Officer in 1998. He holds a B.Eng. and M.Eng. in mechanical engineering specialized in robotics and a Ph.D. in electrical engineering specialized in human factors, all from the University of Montreal. He is pursuing research on the design, realization and evaluation of interactive technologies for training, supervisory control and entertainment purposes. He participated in several research projects involving the use of virtual reality technologies in the forestry, mining, space robotics, heritage and military sectors.

Cosmin Munteanu: Dr. Cosmin Munteanu is a Research Officer with NRC-IIT where he leads the Voice & Multimodal lab and is part of the Mobile HCI lab. For over twelve years, he has conducted research aimed at improving humans' access to and interaction with information-rich media and technologies through natural language. Presently, Cosmin leads several research projects that explore multimodal interaction for mobile devices and mixed reality systems for virtual training. His area of expertise is at the intersection of Automatic Speech Recognition, Natural Language Processing, and Human-Computer Interaction, having published extensively in these areas.

Bruno Emond: Dr. Bruno Emond joined NRC-IIT as a Research Officer in 2001 and holds a B.A. and M.A. in philosophy from the University of Montreal, and a Ph.D. in educational psychology (applied cognitive science) from McGill University (1995). Dr. Emond's current research interests focus on the application of cognitive modeling technology in training simulators, as well as learning and performance in multimedia and broadband e-learning environments.

Irina Kondratova: Dr. Irina Kondratova is a Senior Research Officer and Group Leader at NRC-IIT and a Professional Engineer with the Association of Professional Engineers for the Province of New Brunswick. She is a graduate of the University of New Brunswick where she holds an adjunct professor position at the Engineering faculty. She joined NRC in 2001, when IIT opened in Fredericton, New Brunswick. Dr. Kondratova's current research is in the area of usability and user interface design, including speech-based and multimodal interfaces, mobile interfaces, culturally appropriate user interface design, pervasive health and advanced training applications.

A Multidisciplinary Approach to Enhancing Infantry Training through Immersive Technologies

Fournier H., Lapointe, J.-F., Munteanu C., Emond B., & Kondratova, I.
Institute for Information Technology, National Research Council of Canada
[Helene.Fournier; Jean-Francois.Lapointe; Cosmin.Munteanu; Bruno.Emond; Irina.Kondratova]@nrc-cnrc.gc.ca

CLOSE QUARTER BATTLE TRAINING

Technology requirements gathering

This research project addresses some of the challenges and shortcomings of existing training systems, both physical and virtual, and the specific training needs we have identified through field studies in partnership with the Department of National Defence Canada (DND).

Requirements for learning scenarios for training close quarter battle (CQB) tasks were collected from subject matter experts (SME) and CQB instructors. These requirements were used to develop technology components for training shooting precision and concurrent threat assessment skills, and expanded to include marksmanship skills for military and law enforcement training.

Some of the major challenges faced by today's training methods for CQB include cost, efficiency, effectiveness, and safety. Challenges were further refined in the context of infantry training as part of field studies in partnership with DND as follows:

- High cost: current training resources are not always reusable (e.g. CO_2 cartridges, plastic bullets, broken doors);
- Inefficiency: layouts are not easily reconfigurable (wooden rooms require carpentry work to reconfigure or non-reconfigurable concrete buildings);
- Lack of realism: paper targets are static;
- Safety needs: certain operations require extreme caution (e.g. breaching a door with a C4 explosive charge).

Problems with current training solutions

There is a growing interest in the Canadian Army for using off-the-shelf computer games in training because of the interactivity and engagement they create for the player. However, training simulations and games are designed with different objectives in mind: a game being focused on the entertainment value for the player and a training simulation being focused on the achievement of learning objectives. Roman and Brown (2007) present a comparison table of gamers and trainers' preferences (see Table 1), originally presented by Helsdingen (2006). The table shows important and possibly irreconcilable differences between preferences for these two types of technology users. However, serious games are aiming at bridging this gap, offering engaging situations for the learner, while, at the same time, being designed with learning objectives in mind.

Table 1. Comparison of gamers' and trainers' preferences

Gamer Preferences	Trainer Preferences
Entertainment	Learning Process
Emotion	Structure
Player Control	Learning Goals
Free Play	Instructor Control
Unpredictable Turn of Events	Standardization
Fantasy	Realistic Problems
No Boundaries	Effective and Efficient
Social Interaction	Transfer of Training
Surprise	Validity
Risk	Fidelity
Suspense	
Art and Beauty	

Current approaches and solutions to training also highlight problems such as cost, scalability and reach, with cost being one of the main challenges for the adoption of virtual training solutions (Smith & Steel, 2000). This is particularly relevant for the law enforcement sector, where personnel are located in smaller units, making the procurement of virtual training simulators more challenging. For example, the new FITE initiative of the U.S. Department of Defense has a development budget of US$36 million (Office of Naval Research, 2009). By comparison, the entire annual training budget of the Daytona (Florida) Police Department was US$ 200,000 in 2009, and anticipated to be drastically reduced as reported in 2010 – a situation faced by almost 70% of the police departments in the U.S. (Johnson, 2010). While a commercial version of the FITE system under development is not yet available, it can be expected that such large development costs, together with the required complex infrastructure to support the system, will create challenges for the in-house development and even for the procurement of such solutions by typical law enforcement units, as well as by small or geographically-distributed infantry units.

Within our research project we aimed at addressing cost and maintenance challenges associated with training simulators by developing a system that could be readily acquired, deployed and maintained by a small infantry unit or a police department. The following sections will present the current state of development for our training simulator platform related to task analysis, cognitive modeling, simulated firearm specifications and multimodal interactions. A systematic requirements specification process ensures that the training system is designed to meet the desired level of performance and readiness from soldiers.

A FIELD STUDY OF CANADIAN INFANTRY TRAINING

One issue with live training facilities is their limited ability to simulate different training scenarios for infantry, law enforcement, or public safety. Virtual technologies for enhanced training are becoming increasingly available; however, such technologies are frequently costly and logistically prohibitive. These shortcomings of current training technologies led to the initiation of the collaborative research project between the National Research Council Institute for Information Technology (NRC IIT) and DND Combat Training Centre Tactical School at Canadian Forces Base (CFB) in Gagetown.

Research efforts of the multidisciplinary project team that included researchers in human computer interaction, cognitive modeling, natural language processing and educational technology resulted in development of the flexible mixed-reality training simulator platform.

THE IRET PLATFORM

The purpose of the Immersive Reflexive Engagement Trainer (IRET) is to blend a number of existing technologies to allow soldiers to train simultaneously within virtual and real environments (see Figures 1 and 2). The primary use of the system is to train personnel in the rapid application of judgment to include the application of rules of engagement and the use of force. A secondary purpose of the system is to allow personnel to practice engagement skills with primary and secondary weapons.

The IRET training platform is comprised of a highly customizable set of interconnected technologies (Figure 3) that allows infantry, law enforcement, or public safety personnel to train using a virtual environment in a manner similar to training in a real physical environment.

Within the project we also created a platform for developing and validating novel interaction methods, technologies, and devices that create mixed-reality immersive training systems that are safer, more cost efficient, and more effective for learning than physical environments or purely virtual reality systems.

Figure 1. IRET chamber with three wall projection units

Figure 2. Mobile IRET projection units

Platform architecture

The system is comprised of a training simulator for shooting skills and close quarter battle training. The training simulator includes laser tracking, speech recognition, mobile devices for interacting with the training simulator, and a set of virtual devices to simulate flash-bangs, and explosives. The simulator also includes a physical training environment such as the IRET chamber – a structure with three walls and projectors (Figure 1) or the mobile projection unit (Figure 2).

Decoupling the system components with the use of a message broker allows for an easy way to inter-connect autonomous systems through a publish-and-subscribe mechanism, with message topics as a means to specify device specific or general communication channels (see Figure 3).

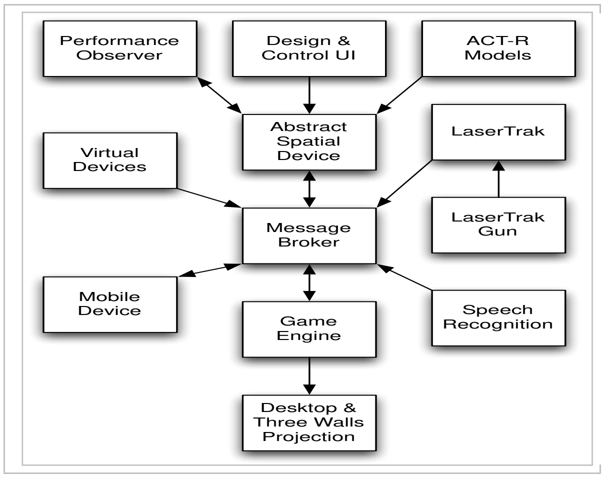

Figure 3. Overall IRET system components

Network-distributed interaction

The interaction with the IRET system is supported by a universal communication protocol implemented through network-based distributed game controls, as well as through the FireWire and STOMP protocols facilitating integration of external components and network-enabled remote training. The software architecture is complemented by a framework for simple configuration of the physical environment using mobile screen projection units. The combination of network-based distributed game engine interaction with the design of a physically flexible training platform allows for easy reconfiguration of the physical layout of the immersive environment to support different training scenarios.

Gaming system

The simulator is made of an immersive room with projection on three walls (or the mobile projection units in other configurations). The game engine used to drive the graphical simulation is based on the C4 engine from Terathon technologies. The purpose of this system is to supply instructors, in a game based-learning environment, with a set of tools to design scenarios and instruction levels as needed. Using the game designer and controller module, an instructor can present a set of floating geometric figures for early shooting skill and judgment learning, or populate the game with opponent avatars controlled by cognitive models (Figure 4).

The user/instructor has an option to animate avatars with a selection from artificial intelligence behaviors. An instructor can choose the game room elements and textures, its contents, drag things around (they simultaneously move within the game). Many object features are controlled through menus, buttons, and pop-up lists.

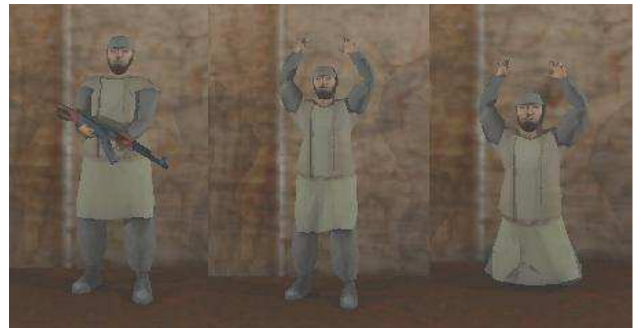

Figure 4. The game environment

Learning objectives

The IRET training system is aimed at substantially improving trainee's skills of physical tactical techniques through intuitive displays and rich feedback, without significant extra time being required for trainees to devote to system-specific training or setting up the system for each scenario, e.g. the users can simply walk up to a training unit and start using it.

Trial observations and meetings with shooting subject matter experts (SME) and close quarter battle (CQB) instructors yielded valuable information to feed into research and development activities and applications. Details were obtained from two different debriefing sessions with SME and CQB instructors regarding the place of simulations in Urban Operations training for continual training and for certain aspects of training—e.g., shot placement and accuracy, reactions for simple entry, for simple baseline training and for non-complex drills, for practicing simple drills such as two-man, four-man entries, and for introducing "barriers" (or obstacles) in simulated drills—e.g., virtual closets, with touch screen interaction to open doors. From these same exchanges, requirements for learning scenarios for training CQB tasks were derived and later validated by SME and CQB instructors and field observations; details that were applied directly to the development of technology components for training shooting precision and concurrent threat assessment skills. Requirements were further explored and eventually expanded through review and discussion to include marksmanship skills for military and law enforcement training.

Voice and multimodal interaction

The IRET training context is extremely noisy; as such, it presents significant, environmentally-generated challenges to the efficacy of speech-based interaction. Ambient noises – such as gunfire, flash bangs, and the simultaneous shouts of multiple soldiers – will, very likely, negatively impact achievable speech recognition rates. Additionally, little is currently known about the comparative impact of different speaking styles – e.g., shouting versus talking versus whispering; most speech recognition evaluation has focused solely on normal speech (i.e., talking). Furthermore, in the case of the IRET simulator, soldiers will be exposed to various stress-inducing situations which are also likely to negatively impact speech patterns. Ambient noise, speaking style, and stress-induced speech variations are just some of the challenges presented by the IRET simulator.

Adding speech recognition capabilities allows for a more natural interaction with an immersive system. However, due to the adverse conditions mentioned previously, the accuracy of speech recognition can be severely degraded. One possible solution that facilitates speech-based interactions in unfavorable conditions is to enable a more complete multimodal interaction – exploiting complementary modalities has been shown to increase the success rate of specific actions (Oviatt, 2003). Based on this premise, we developed the Multimodal Interaction (MINT) component of the IRET system.

MINT has been designed to support customizable training scenarios for a wide range of domains; it consists of a set of interconnected technologies: speech interface, simulated electronic weapons, and touch interfaces for dynamic game interaction, all interfacing with the implementation of a flexible and scalable framework that controls the immersive environment (Figure 5).

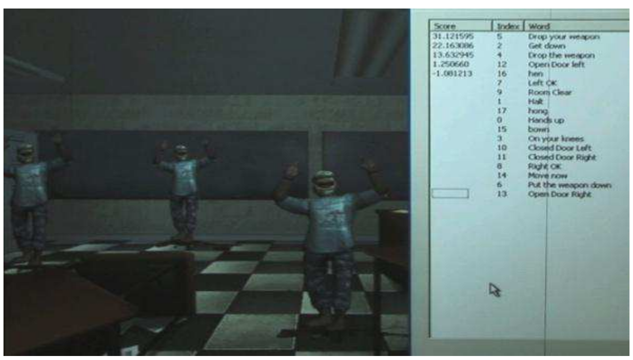

Figure 5. Multimodal components in IRET system

The speech interface

Trainees can interact with the gaming system through spoken commands. The speech processing unit is implemented as a software interface for connecting, controlling, and configuring speech recognition and language processing for natural interaction with virtual characters through voice commands. The key component of the speech interface is a back-end implementation of flexible and robust grammar-based processing of speech under stress, allowing the connection with intuitive grammar editors. In order to accommodate various training needs, both in terms of suitability for a particular setting and in terms of budget constraints, the MINT components can be configured to run with several commercially-available speech recognition systems, by implementing connectivity between a third-party speech engine, grammar-based processing, and the game platform.

Artifact-based interaction

The flexible design of the MINT components allows trainees to interact with the gaming system through several modalities, such as gestures or weapons. In the current implementation of the IRET system we have built and evaluated several electronic devices simulating lethal and non-lethal weapons that interact with the immersive environment, as determined by the needs analysis conducted with our partner (DND) through field studies and observations. For this, we have designed and developed functional prototypes of several simulated weapons, such as flashbang and explosive door breacher. These simulated weapons communicate with the gaming system and allow for a rich interaction within the training environment. This is accomplished through the implementation of a hardware-software architecture that controls communication between devices and the gaming platform and supports further enhancement of the immersive interaction through physical embodiments of virtual elements.

Touch-based interaction

Beside trainees, instructors are also users of a training system. As such, appropriate interactive components need to be provided to enhance the overall realism of the training scenarios. In contrast to traditional training methods, an immersive gaming environment allows for richer and more engaging interactions. To address this, we developed a tablet-based control interface for trainers/instructors that facilitates real-time manipulation of training scenarios (see Figure 6).


Figure 6. Trainer control interface

The trainer control unit is implemented as a touch-based interface serving as a teaching tool that allows instructors to animate game characters in response to trainees' actions. All the animations and actions that the game can perform can be initiated through this trainer control interface. Moreover, game actions that are initiated by trainees through spoken commands can be overridden by trainers (e.g. when the speech recognition system did not accurately process the command) or reversed by trainers, e.g. to create “on-the-fly” training situations that test the responsiveness and judgment of the trainees, such as, for example, introducing non-compliant behavior for the virtual game characters.

Laser weapons

In order to simulate the training with real weapons, we developed laser rifles that have the same shape and weight as the real rifles used for combat, namely the Colt C7A2. These laser rifles are made of mockup rifles equipped with a laser source and a battery pack for the power supply. The laser rifle is illustrated in Figure 7 below.

Figure 7. The laser rifle

The required electronics have been made to allow firing short pulses of laser beams each time the trigger is pulled. This way, trainees can shoot up to about 13 virtual round/s, which is the firing rate of this firearm when shooting in fully automatic mode.

More than one rifle can be tracked simultaneously by the vision-based system used to track the laser spots. It is therefore possible to identify not only the number of shots but also the source of the shots. This capability is a key component of the training simulator, allowing for simultaneous training of several trainees while keeping track of individual performance.

Cognitive modeling

Cognitive modeling, as a methodology, has its roots in cognitive architecture research. Cognitive architectures are general in the sense that they claim to be universally capable of modeling cognitive activity. Cognitive architectures are relatively complete proposals about the structure of human cognition. A cognitive architecture provides the resources for developing models. These resources take the form of a set of specifications regarding functional invariants (Pylyshyn, 1984) related to knowledge representation, knowledge processing, memory, perception, and motor actions. Some examples of cognitive architectures are SOAR (Newell, 1990), ACT-R (Anderson et al., 2004; Anderson & Lebiere, 1998), EPIC (Kieras & Meyer, 1997), and CI (Kintsch, 1998). Each of these architectures has its strength and was initially developed with some intended modeling purpose. ACT-R is mainly focused on problem solving and memory as well as vision and motor control (Byrne & Anderson, 2001), SOAR on problem solving and learning, EPIC on multiple task performance, and CI on text comprehension. The cognitive modeling methodology is mostly an iterative methodology similar to the learning cycle in HCI research (Olson & Olson, 1997) and computational psycholinguistics (Dijkstra & De Smedt, 1996). Simulations, as a method a scientific inquiry, can best be advanced and tested by the concurrent and complementary use of empirical methods (Goldspink, 2000).

Associated to the task analysis, the project also conducted a cognitive modeling activity to develop constructive simulations of the CQB skills. Constructive simulations are key elements in the development of training simulators. They can be used to help in the acquisition process, as a foundation for the development of synthetic adversaries, as a means to detail the skills to be acquired in a training simulator, or even to study the transfer of agent skills. A broader access to game engines as well as the emergence of new or improved cognitive architectures has allowed the development of many simulation systems of military operations on urban terrains.

The cognitive modeling research activity within the IRET project has two principal objectives: a) develop high-fidelity cognitive modeling technology to be embedded as artificial intelligent agents in an immersive combat game; and b) develop detailed performance and learning models of the learners to support instructions. The next section will discuss the current state of development in achieving the first object.

There are many definitions of what an agent is but the following characteristics seem to describe adequately what being an agent means. An agent is an identifiable, discrete individual. It is autonomous and self-directed (goal driven); it is situated, living in an environment with which it interacts with other agents (having perceptual, motor, and communication capacities); and it is flexible, having the ability to learn and adapt its behaviors based on experience. Agent-based modeling is divided in two communities, one focused on large numbers of relatively simple and highly-interactive agents; and the other one focused on a smaller number of agents with more complex internal structures. Our current research falls into the second category, and uses the ACT-R cognitive architecture as a means to develop agents, which has modules to implement goal driven behavior, perceptual and motor capabilities, as well as learning mechanisms.

For the purpose of this paper, cognitive models and agents will be considered synonymous. However, because the modeling approach is based on the ACT-R cognitive architecture, when a reference is made to a cognitive model, the internal structure of the model is the point of interest, such as the perceptual and motor modules, or the declarative and procedural memory modules. On the other hand, when the point of interest is not the internal structure but the individual and discrete nature of an entity, then the term agent will be used.

Figure 8 presents a flow diagram of the cognitive modeling methodology spanning from task analysis to model verification and validation. Processes are represented as ellipses and products as rectangles. A first distinctive feature of the approach is the development of an environment model in parallel to the cognitive model. The environment model is a piece of software with which a cognitive model or a human user can interact. Only relevant characteristics of the environment for the tasks that need to be performed are included in the environment model. The same executable environment model can be used to collect data on human performance, and provides the perceptual and motor environment for the cognitive model. The figure also shows that model verification proceeds by comparing simulated performance data to the task formal model, while model validation proceeds by comparing simulated to human performance data.

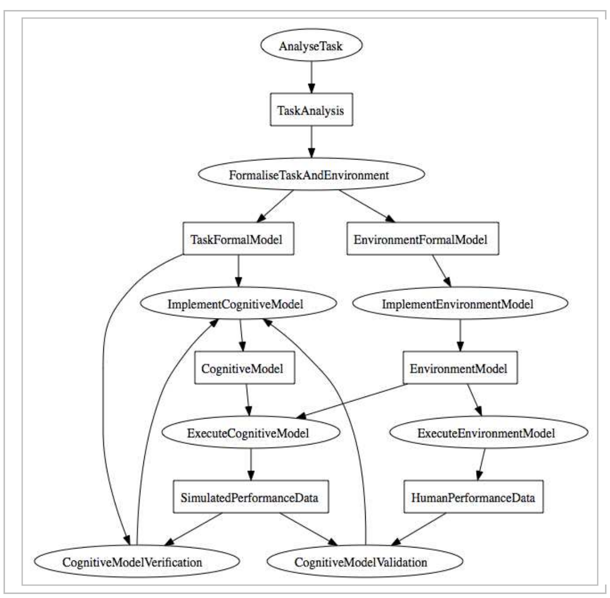

Figure 8. Cognitive and environment modeling process

The interdependency between simulated agents and the environment in which they interact puts forward the need for cognitive model specifications to include both cognitive elements such as perceptual and motor skills, as well as environment affordances. A constructive simulation needs to identify the high-level primitive perceptual representations and motor actions essential for a cognitive model to interact with a simulated environment. As depicted in Figure 9, the flow of information between a device and a cognitive architecture requires an interface to bind internal device data to perceptual presentations, and motor actions to device inputs. These primitives constitute the first set of modeling requirements.

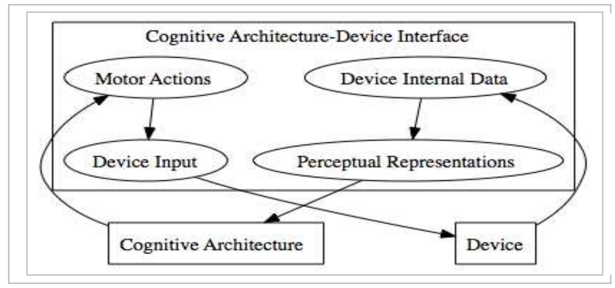

Figure 9. Information flow

The intermediate layer between a cognitive architecture and devices, such as a desktop application or a game engine, can be described by functions transforming internal device data into high-level perceptual constructs feeding in the cognitive model perceptual modules. In the same manner, motor actions get executed in the external device by translating high-level action representations in the cognitive model into device input.

Initiative based tactics are driven by the actions and initiative of the individual soldiers. Proper actions must conform to the doctrine and fundamentals of close quarter battle (CQB), but the actions success is highly dependent on the application of skills directed by the challenges of the immediate and specific conditions of a CQB situation. Communication and coordination with teammates, efficient body movements, as well as rapid threats assessment from environmental cues are important building blocks of initiative-based tactics skills.

Prior research in CQB tasks analysis and cognitive modeling applications provide an initial identification of key perceptual and motor primitives. Table 2 summarizes some of these primitives. The table is divided into perceptual and motor modalities. Most of the categories and labels should be relatively easy to understand, such as location and end-points (defined in an egocentric spatial coordinate system), volume, and type. The people category however identifies environmental affordances that are crucial to the assessment of a threat level. Acquired-visual-object and weapon-target for example are the respective projections of the line of sight and weapon pointing direction onto agents in the room. Weapon readiness and potency are also other perceptual factors in threat assessment. A person can also exhibit composition of course and heading variations to produce different kinds of body motion such as steering (aligned course and heading); canted (fix alignment offset between course and heading), oblique (constant heading position), and scanning (free heading movement from the course).

All of the primitive perceptual, cognitive and motor elements presented in Table 2 can be physically measured with some level of accuracy using sensors located in a training room or worn by the participants. Sensor data (ex. tracking body movements) associated with trainee performance assessments by experts would allow a refined performance analysis that would create mathematical behavior models based on real-world actions. During the requirement specifications stage, one of the purposes of a cognitive model is to provide an initial mapping between simulated physical measurements, model performance, and cognitive operations, which are then represented as a set of production rules; within these production rules, perceptual, cognitive, and motor operations can be linked to observable objects and behaviors.

Table 2. Perceptual, cognitive and motor constructs required to operate in a CQB situation.

Perception Audition	
Verbal messages	Location; Volume; Sender ; Content
Weapon fire	Location ; Volume ; Type
Ricochets	Location ; Volume ; Type
Flash bang	Location; Volume
Footsteps	Location ; Volume ; Direction
Perception Visual	
Non-verbal messages	Sender; Content
Walls	End-points ;
Corners	Location ;
Pathways	End-points ;
Doors	End-points; Hinges-location; Open-state;
Weapons	Location ; Type
Objects	Location ; Type
People	Location ; Type ; Speed ; Course ; Heading ; Acquired-visual-object; With-weapon; Weapon-potency; Weapon-orientation; Weapon-readiness; Weapon-target
Motor Communication	
Speech	Receiver; Content; Volume
Non-verbal messages	Receiver; Content
Motor Body	
Weapon handling	Type; Trigger-arm&hand; Readiness ; Orientation; Pull-Trigger; Throw ;
Body displacement	Course ; Heading Speed ; Modality
Body rotation	Heading ; Speed

FIELD STUDY and EVALUATION

Preliminary evaluation

IRET was first used in the context of Urban Operations training for simple baseline training in practicing shooting precision (e.g., shot placement and accuracy), threat assessment, for the purpose of practicing simple drills.

IRET demonstrations and trials with SME and CQB instructors also provided researchers with an opportunity to collect data which included notes from observations during live IRET testing exercises and demonstration sessions, commentary from SME and CQB instructors while interacting with the system and its components, as well information from debriefing sessions held with participants. The information gleaned from these punctual exchanges revealed the need for learning scenarios to be expanded through better customization of game parameters (e.g., hits, penalties, times), for improved capacity for tracking teams playing with multiple laser detection, and for better scenarios for discriminating target depth and perceived distance in the context of developing shooting precision and threat assessment skills. Modifications to the game and IRET system were commented on and validated by SME and CQB instructors during subsequent field trials and demonstrations.

Evaluation of multimodal components

The design of the multimodal components (MINT) within the IRET platform has been informed by the observations carried out during field training by infantry trainees at the Canadian Forces Base (CFB) in Gagetown. The field study was conducted over a period of two weeks, one during spring training and one during fall training, covering all stages of basic infantry training. The requirement gathering was focused on identifying aspects of current training procedures that can be performed under more realistic, more cost effective, and safer conditions in a lab-based environment. Among such conditions we have identified the use of area weapons, explosive-based breachers, and engagement of enemy combatants (targets) as the first candidates for our prototype system.

The first iterations of the IRET-MINT components were validated through a pilot study in a lab-based setting lasting three days, with volunteer participants recruited from among the course instructors at CFB Gagetown (for a total of six participants). The participants conducted simulated training scenarios in a large laboratory using the IRET equipment, aiming to replicate the same scenarios as for the regular infantry training. Data was collected from this pilot evaluation in the form of direct observations, video recordings, and semi-structured interviews with the participants.

The analysis of data collected from the pilot evaluations was focused on identifying improvements required for the IRET-MINT components, both at implementation level and at design / interface level. The evaluation confirmed the value of trainers to the training environment. To address this, we have augmented the interaction to allow trainers to manipulate the game scenarios dynamically through a tablet-based touch interface.

The revised version of the IRET-MINT component was fully evaluated through a complete lab-based study over five days. The study was conducted in a similar fashion as the pilot evaluation. Five instructors from CFB Gagetown participated voluntarily in the evaluation. While the data collected from the final evaluation is currently being analyzed, preliminary findings from the study suggest that:

- The virtual training system is well received by its intended users.
- Speech-based interaction is critical in ensuring the realism of the interaction with the virtual environment.
- The accuracy of automatic speech recognition can vary significantly across training conditions, but can be easily compensated for by using complementary modalities, such as the tablet-based trainer interface.
- The ability to quickly configure the physical layout of the training environment is an important feature that significantly reduces the down-time between training sessions.
- The simulated lethal and non-lethal devices that interact with the gaming environment (e.g. flashbang, door breacher) were seen by participants as one of the most important components of the mixed reality training system.
- The touch-based tablet interface for trainers enabled trainers to quickly customize and control the scenarios, as well as to react to trainees' mistakes in a timely fashion and was very well received by trainers.
- Although our system currently does not have a full implementation of after action review (only video cameras recording the users' actions), the participants indicated that the ability to play back the entire interaction with the gaming environment would be a valuable educational tool.

OUTCOMES AND LESSONS LEARNED

IRET trial observations and validation exercises with shooting SMEs and CQB instructors yielded valuable information on the significance of simulations in Urban Operations training and on important parameters the instructors would like to control in the training system:
1) shape of the room;
2) types of drills;
3) target distance;
4) features of the system (e.g., random generation of scenarios, ability to turn on/off feedback option, etc).

Observations and trials of IRET technology also revealed that SMEs and CQB instructors preferred a blended approach for basic to advanced skills training, including target selection and shooting games for warm up exercises and skills maintenance, and mixed approaches for complex entry drills, including mobile displays and VR displays with avatars and wall projections, mobile capability for changing the game scenario in real time, real objects such as chairs and tables as obstacles in the scene, and the IRET chamber for breaching and entry drills more specifically. Trial participants underscored the advantages of using IRET component technologies for training marksmanship skills both in a military or law enforcement context. This last validation exercise reinforced the potential for expansions of the current applications to include: better customization of game parameters such as points earned and penalties, integration of multiple lasers detection within the game for team playing, addition of target depth and perceived distance, development of a shooting assessment tool based on the target selection and shooting game cognitive model, development of weapon movement, and eye tracking measurements to capture in detail shooting and threats assessment skills, as well as real-time measurement of stress physiological indicators.

Technology outcomes

As part of our partnership with the Combat Training Center Tactical School at CFB Gagetown within the collaboration with Department of National Defence (DND) Canada, the Multimodal Interaction platform of the IRET project has been customized and implemented to meet the specific needs of our partner. The focus of this work has been to investigate natural and intuitive interactions for effective use within a combat training simulator. The customizations based on users needs, as determined through field studies and lab evaluations, resulted in several research and technological outcomes:

- Identification of key human factors for interacting by voice with a combat training simulator.
- Research on Automatic speech recognition (ASR) performance issues in real-life training scenarios.
- Methodologies for the successful adaptation and use of ASR within training environments, taking into

account the specific human factors and performance limitations.

- Remote-input technologies to complement speech interactions and allow realistic use of simulated weapons and non-lethal weapons.
- The ability to adjust precisely the aiming of the laser beam on the laser rifles with the addition of a mechanical adjustment system using screws.

CONCLUSIONS

The results of the research conducted as part of this project lead to important recommendations regarding the feasibility of having an immersive (3-wall) multi-user system where the users interact with input devices emitting laser lights, with simulated area weapons and breachers, and with virtual environments through speech and touch. Field trials and demonstrations with our partner (DND) resulted in the development of some alternatives to the IRET chamber, including mobile screens equipped with a laser shooting interface to allow for quick reconfiguration of training rooms for developing shooting skills, as well as mobile, tablet-based, touch interfaces allowing trainers to dynamically control the game interaction. Studies on the design and interactions with such units could be an interesting research path for a follow-up project with either military or law enforcement end users. Possible extensions of the current version of the IRET mixed reality training platform need to address the following issues highlighted by our current research and evaluation: better customization of game parameters such as points earned and penalties, integration of multiple lasers detection in to the game for team playing, addition of target depth and perceived distance, development of a shooting assessment tool based on the target selection and shooting game cognitive model, development of weapon movement, and eye tracking measurements to capture in detail shooting and threats assessment skills, as well as real-time measurement of stress physiological indicators. Furthermore, subsequent developments of the virtual training system should place increased emphasis on several aspects related to human factors of interacting with immersive environments, such as the accuracy of automatic speech recognition under adverse and high-stress conditions, the interaction with physical embodiments of virtual elements, the usability and usefulness of mobile trainer controls, and the ability to remotely supervise training.

ACKNOWLEDGEMENTS

We would like to acknowledge the contributions of the project collaborators and evaluation participants from the DND Combat Training Center in Gagetown. We also thank our colleagues from the Institute for Information Technology of the National Research Council Canada who has been involved in research, design, development and testing related to IRET system, especially to research staff that significantly contributed to project innovations: Rene Richard, Cedric Martin, Muhammad Al-Digeil, Brian McKinnon, Jared Vawter, Haozheng Li, Ilia Goldfarb, Nathan Langton, Yasaman Sanaee, Candace Doucet, and Cajetan Bouchard.

REFERENCES

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. Psychologica Review, 11(4), 1036-1060.

Anderson, J. R., & Lebiere, C. (1998). The Atomic Components of Thought. Mahwah, NJ: Lawrence Erlbaum Associates.

Byrne, M. D., & Anderson, J. R. (2001). Serial modules in parallel: The psychological refractory period and perfect time-sharing. Psychological Review, 108, 847-869.

Dijkstra, A. F. J., & De Smedt, K. J. M. J. (1996). Computer models in psycholinguistics. In A. F. J. Dijkstra & K. J. M. J. De Smedt (Eds.), Computational psycholinguistics: AI and connectionist models of human language processing (pp. 3-23). London: Taylor & Francis.

Goldspink, C. (2000). Modelling social systems as complex: towards a social simulation meta-model. Journal of Artificial Societies and Social Simulation, 3(2). Retrieved from http://jasss.soc.surrey.ac.uk/3/2/1.html.

Helsdingen, A. (2006). *Games for Training*. Paper presented at the VV&A Methods, Defense Simulation and Training Conference.

Kieras, D. E., & Meyer, D. E. (1997). An overview of the EPIC architecture for cognition and performance with application to human-computer interaction. Human-Computer Interaction, 12, 391-438.

Kintsch, W. (1998). Comprehension. Cambridge: Cambridge University press.

National_Research_Council. (2002). *Modeling and Simulation in Manufacturing and Defense Acquisition: Pathways to Success*. Washington, DC: National Academy Press.

Newell, A. (1990). Unified theories of Cognition. Cambridge, MA: Harvard University Press.

Office of Naval Research. (2009). *Office of Naval Research Selected to Lead Technical Management of Vital New Joint Training Program.* (Media Release, January 30, 2009, retrieved from http://www.onr.navy.mil/Media-Center/Press-Releases/2009/).

Olson, G. M., & Olson, J. S. (1997). Research on computer supported cooperative work Handbook of Human computer interaction (pp. 1433-1457). Amsterdam: Elsevier.

Oviatt, S. (2003). Advances in Robust Multimodal Interface Design. *IEEE Comput. Graph. Appl. 23(5),* (pp. 62-68).

Pylyshyn, Z. W. (1984). Computation and cognition: toward a foundation for cognitive science. Cambridge, MA: MIT Press.

Roman, P. A., & Brown, D. (2007). Note to file - Constructive simulation versus serious games for the army: a Canadian case study. *Canadian Army Journal, 10*(3), 80-88.

Smith, J.D. & Steel, J. (2000). The use of virtual simulation for dismounted infantry training. *Proceedings of the 2nd NATO Modelling and Simulation Conference* (pp. 15:1-15:9).